ANASIEZE IKENNA – CLOUD & AI SOULTIONS ENGINEER

Project: Real Time Carbon Storage Monitoring Pipeline with BigQuery

Overview

This project demonstrates how I designed and deployed a real time Carbon Capture and Storage (CCS) data streaming pipeline on Google Cloud. The system ingests live sensor data from CO_2 injection wells, processes it continuously, and stores it in BigQuery for immediate querying and analysis.

Rather than relying on batch uploads or manual reporting, this architecture supports continuous CCS sensor data, making it suitable for use cases such as injection monitoring, storage integrity oversight, and operational analytics for safe long term carbon storage.

Problem Statement

Oil field engineers supporting CCS operations needed a solution to:

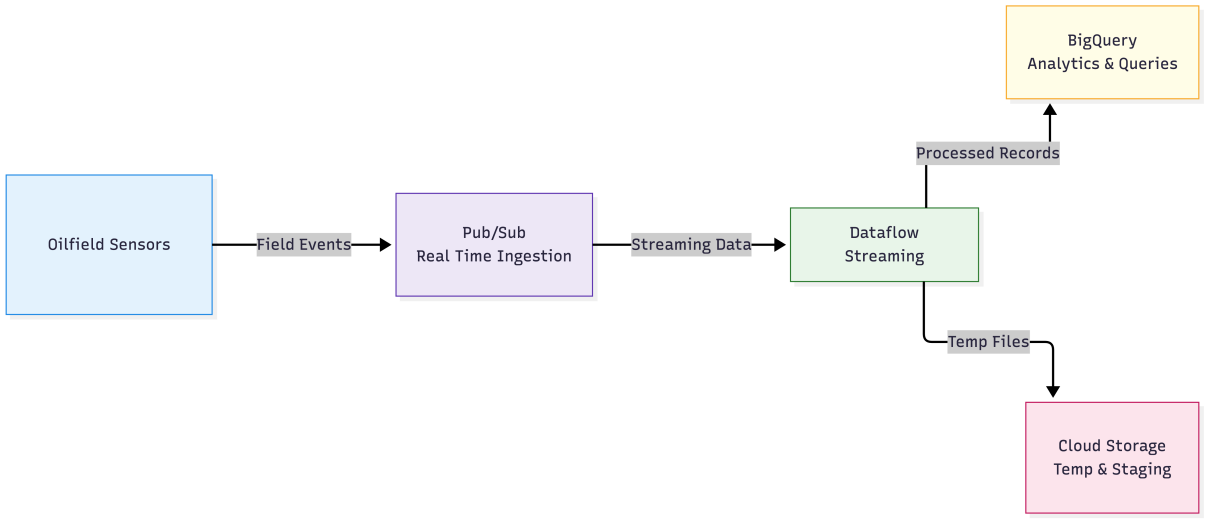
+ Ingest real time sensor data from CO_2 injection wells
+ Process continuous pressure and temperature data reliably at scale
+ Store CCS sensor data in an analytics ready data warehouse
+ Validate incoming field data as it arrives

Goal: Responsible for designing and implementing the end to end CCS sensor data streaming pipeline using Google Cloud managed services.

Tech Stack

+ Cloud Storage – temporary and staging storage for Dataflow job execution
+ Pub/Sub – real time ingestion of oilfield sensor data
+ Dataflow – continuous stream processing of field data
+ BigQuery – analytics storage for querying and validation

Architecture Overview



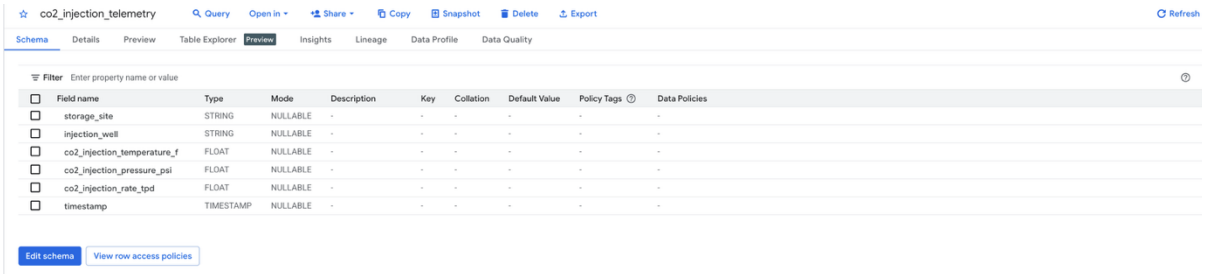
Flow of Data

1. CO_2 injection well sensors publish sensor data events
2. Pub/Sub ingests CCS sensor data in real time
3. Dataflow processes the sensor data stream continuously
4. BigQuery stores processed records for analytics and reporting
5. SQL queries are used to validate data and analyze injection conditions

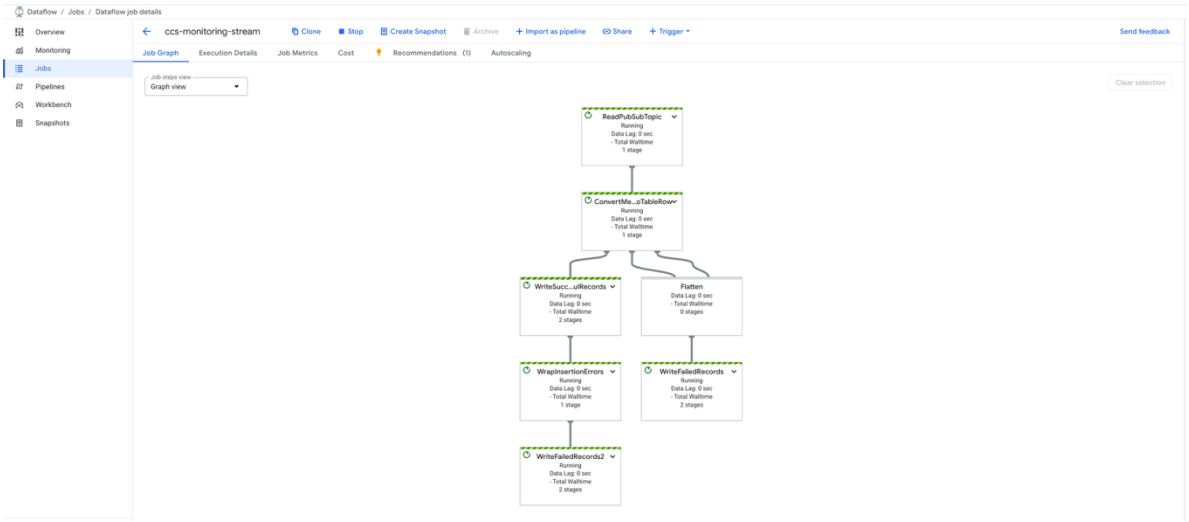
This architecture is serverless, scalable, and aligned with modern digital oilfield platforms.

Deployment Steps

+ **Provisioned Cloud Storage for Dataflow execution job**
+ **Created a BigQuery dataset and table for streaming ingestion**

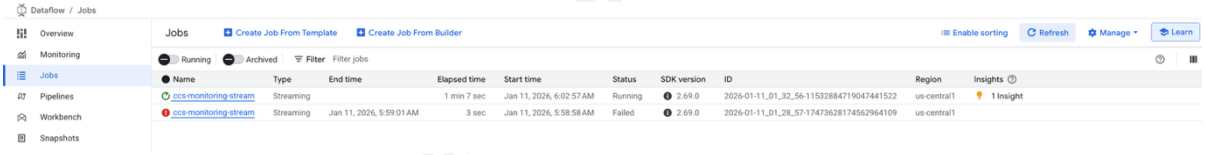


+ **Configured Pub/Sub for real time oilfield sensor data ingestion**



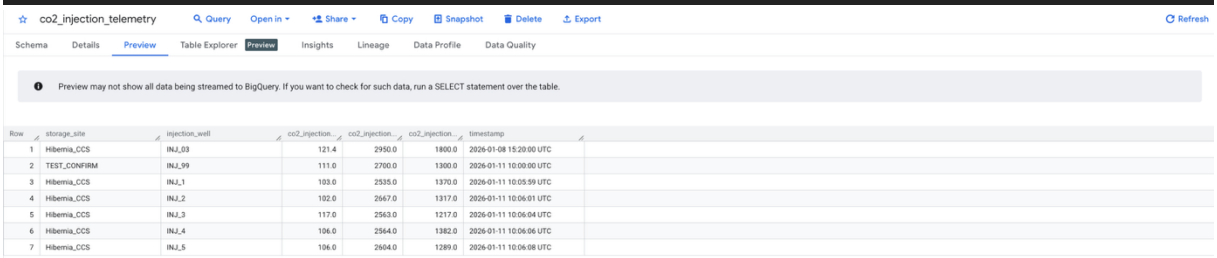
+ **Deployed a Dataflow streaming pipeline from Pub/Sub to BigQuery**

```
JOB_ID: 2026-01-08_19_05_35-14579305350506963036
NAME: iotflow
TYPE: Streaming
CREATION_TIME: 2026-01-09 03:05:36
STATE: Running
REGION: us-central1
```



+ **Published sensor data events and validated live ingestion using BigQuery queries**

```
do
  gcloud pubsub topics publish $TOPIC_NAME \
    --message="{
      \"data\": {
        \"storage_site\": \"Hibernia_CCS\",
        \"injection_well\": \"INJ_$((RANDOM % 10))\",
        \"co2_injection_temperature_f\": $((90 + RANDOM % 60)),
        \"co2_injection_pressure_psi\": $((2000 + RANDOM % 1500)),
        \"co2_injection_rate_tpd\": $((500 + RANDOM % 2000)),
        \"timestamp\": \"$(date -u +%Y-%m-%dT%H:%M:%SZ)\"
      }
    }"
done
messageIds:
- '17725029986365256'
messageIds:
```



Each component was created in the correct region with required APIs enabled to reflect production best practices.

Outcome

- A fully operational real time oilfield sensor data streaming pipeline.
- Live sensor data flowing continuously from Pub/Sub into BigQuery.
- A reproducible, cloud native analytics architecture suitable for oil and gas operations.

This project shows that I can design, deploy, and operate real time oilfield data pipelines, troubleshoot distributed cloud systems independently, and think end to end like a cloud engineer supporting energy and production environments.